



1

SEQUENCE LISTING

<110> EVANS, RONALD M.

<120> XENOBIOTIC COMPOUND MODULATED EXPRESSION SYSTEMS AND
USES THEREFOR

<130> 088802-5211

<140> 09/840,008

<141> 2001-04-20

<150> 09/227,718

<151> 1999-01-08

<150> 09/458,366

<151> 1999-12-09

<150> 09/005,286

<151> 1998-01-09

<160> 44

<170> PatentIn Ver. 2.1

<210> 1

<211> 2068

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (583)..(1887)

<400> 1

ggcagcagga gatctaggtt caaattaatg ttgcccttag tggtaaagga cagagaccct 60
cagactgatg aaatgcgctc agaattactt agacaaagcg gatatttgcc actctcttcc 120
ccttttcctg tgtttttgta gtgaagagac ctgaaagaaa aaagtaggga gaacataatg 180
agaacaaata cggtaatctc ttcatttgct agttcaagtg ctggacttgg gacttaggag 240
gggcaatgga gccgcttagt gcctacatct gacttggact gaaatatagg tgagagacaa 300
gattgtctca tatccgggga aatcataacc tatgactagg acgggaagag gaagcactgc 360
ctttacttca gtgggaatct cggcctcagc ctgcaagcca agtggttcaca gtgagaaaag 420
caagagaata agctaatact cctgtcctga acaaggcagc ggctccttgg taaagctact 480
ccttgatcga tcctttgcac cggattgttc aaagtggacc ccaggggaga agtcggagca 540
aagaacttac caccaagcag tccaagaggc ccagaagcaa ac ctg gag gtg aga 594
Met Glu Val Arg

ccc	aaa	gaa	agc	tgg	aac	cat	gct	gac	ttt	gta	cac	tgt	gag	gac	aca	642
Pro	Lys	Glu	Ser	Trp	Asn	His	Ala	Asp	Phe	Val	His	Cys	Glu	Asp	Thr	
5					10				15						20	
gag	tct	gtt	cct	gga	aag	ccc	agt	gtc	aac	gca	gat	gag	gaa	gtc	gga	690
Glu	Ser	Val	Pro	Gly	Lys	Pro	Ser	Val	Asn	Ala	Asp	Glu	Glu	Val	Gly	
				25					30						35	
ggg	ccc	caa	atc	tgc	cgt	gta	tgt	ggg	gac	aag	gcc	act	ggc	tat	cac	738
Gly	Pro	Gln	Ile	Cys	Arg	Val	Cys	Gly	Asp	Lys	Ala	Thr	Gly	Tyr	His	
			40					45					50			
ttc	aat	gtc	atg	aca	tgt	gaa	gga	tgc	aag	ggc	ttt	ttc	agg	agg	gcc	786
Phe	Asn	Val	Met	Thr	Cys	Glu	Gly	Cys	Lys	Gly	Phe	Phe	Arg	Arg	Ala	
		55					60					65				
atg	aaa	cgc	aac	gcc	cgg	ctg	agg	tgc	ccc	ttc	cgg	aag	ggc	gcc	tgc	834
Met	Lys	Arg	Asn	Ala	Arg	Leu	Arg	Cys	Pro	Phe	Arg	Lys	Gly	Ala	Cys	
	70					75					80					
gag	atc	acc	cgg	aag	acc	cgg	cga	cag	tgc	cag	gcc	tgc	cgc	ctg	cgc	882
Glu	Ile	Thr	Arg	Lys	Thr	Arg	Arg	Gln	Cys	Gln	Ala	Cys	Arg	Leu	Arg	
85						90				95					100	
aag	tgc	ctg	gag	agc	ggc	atg	aag	aag	gag	atg	atc	atg	tcc	gac	gag	930
Lys	Cys	Leu	Glu	Ser	Gly	Met	Lys	Lys	Glu	Met	Ile	Met	Ser	Asp	Glu	
				105					110					115		
gcc	gtg	gag	gag	agg	cgg	gcc	ttg	atc	aag	cgg	aag	aaa	agt	gaa	cgg	978
Ala	Val	Glu	Glu	Arg	Arg	Ala	Leu	Ile	Lys	Arg	Lys	Lys	Ser	Glu	Arg	
			120					125					130			
aca	ggg	act	cag	cca	ctg	gga	gtg	cag	ggg	ctg	aca	gag	gag	cag	cgg	1026
Thr	Gly	Thr	Gln	Pro	Leu	Gly	Val	Gln	Gly	Leu	Thr	Glu	Glu	Gln	Arg	
			135				140					145				
atg	atg	atc	agg	gag	ctg	atg	gac	gct	cag	atg	aaa	acc	ttt	gac	act	1074
Met	Met	Ile	Arg	Glu	Leu	Met	Asp	Ala	Gln	Met	Lys	Thr	Phe	Asp	Thr	
	150					155					160					
acc	ttc	tcc	cat	ttc	aag	aat	ttc	cgg	ctg	cca	ggg	gtg	ctt	agc	agt	1122
Thr	Phe	Ser	His	Phe	Lys	Asn	Phe	Arg	Leu	Pro	Gly	Val	Leu	Ser	Ser	
165					170					175					180	
ggc	tgc	gag	ttg	cca	gag	tct	ctg	cag	gcc	cca	tcg	agg	gaa	gaa	gct	1170
Gly	Cys	Glu	Leu	Pro	Glu	Ser	Leu	Gln	Ala	Pro	Ser	Arg	Glu	Glu	Ala	
				185					190						195	
gcc	aag	tgg	agc	cag	gtc	cgg	aaa	gat	ctg	tgc	tct	ttg	aag	gtc	tct	1218
Ala	Lys	Trp	Ser	Gln	Val	Arg	Lys	Asp	Leu	Cys	Ser	Leu	Lys	Val	Ser	
			200					205					210			
ctg	cag	ctg	cgg	ggg	gag	gat	ggc	agt	gtc	tgg	aac	tac	aaa	ccc	cca	1266
Leu	Gln	Leu	Arg	Gly	Glu	Asp	Gly	Ser	Val	Trp	Asn	Tyr	Lys	Pro	Pro	
		215					220					225				

gcc gac agt ggc ggg aaa gag atc ttc tcc ctg ctg ccc cac atg gct	1314
Ala Asp Ser Gly Gly Lys Glu Ile Phe Ser Leu Leu Pro His Met Ala	
230 235 240	
gac atg tca acc tac atg ttc aaa ggc atc atc agc ttt gcc aaa gtc	1362
Asp Met Ser Thr Tyr Met Phe Lys Gly Ile Ile Ser Phe Ala Lys Val	
245 250 255 260	
atc tcc tac ttc agg gac ttg ccc atc gag gac cag atc tcc ctg ctg	1410
Ile Ser Tyr Phe Arg Asp Leu Pro Ile Glu Asp Gln Ile Ser Leu Leu	
265 270 275	
aag ggg gcc gct ttc gag ctg tgt caa ctg aga ttc aac aca gtg ttc	1458
Lys Gly Ala Ala Phe Glu Leu Cys Gln Leu Arg Phe Asn Thr Val Phe	
280 285 290	
aac gcg gag act gga acc tgg gag tgt ggc cgg ctg tcc tac tgc ttg	1506
Asn Ala Glu Thr Gly Thr Trp Glu Cys Gly Arg Leu Ser Tyr Cys Leu	
295 300 305	
gaa gac act gca ggt ggc ttc cag caa ctt cta ctg gag ccc atg ctg	1554
Glu Asp Thr Ala Gly Gly Phe Gln Gln Leu Leu Glu Pro Met Leu	
310 315 320	
aaa ttc cac tac atg ctg aag aag ctg cag ctg cat gag gag gag tat	1602
Lys Phe His Tyr Met Leu Lys Lys Leu Gln Leu His Glu Glu Glu Tyr	
325 330 335 340	
gtg ctg atg cag gcc atc tcc ctc ttc tcc cca gac cgc cca ggt gtg	1650
Val Leu Met Gln Ala Ile Ser Leu Phe Ser Pro Asp Arg Pro Gly Val	
345 350 355	
ctg cag cac cgc gtg gtg gac cag ctg cag gag caa ttc gcc att act	1698
Leu Gln His Arg Val Val Asp Gln Leu Gln Glu Gln Phe Ala Ile Thr	
360 365 370	
ctg aag tcc tac att gaa tgc aat cgg ccc cag cct gct cat agg ttc	1746
Leu Lys Ser Tyr Ile Glu Cys Asn Arg Pro Gln Pro Ala His Arg Phe	
375 380 385	
ttg ttc ctg aag atc atg gct atg ctc acc gag ctc cgc agc atc aat	1794
Leu Phe Leu Lys Ile Met Ala Met Leu Thr Glu Leu Arg Ser Ile Asn	
390 395 400	
gct cag cac acc cag cgg ctg ctg cgc atc cag gac ata cac ccc ttt	1842
Ala Gln His Thr Gln Arg Leu Leu Arg Ile Gln Asp Ile His Pro Phe	
405 410 415 420	
gct acg ccc ctc atg cag gag ttg ttc ggt atc aca ggt agc tga	1887
Ala Thr Pro Leu Met Gln Glu Leu Phe Gly Ile Thr Gly Ser	
425 430	
gtggctgtcc ttgggtgaca cctccgagag gtagttagac ccagagccct ctgagtcgcc	1947
actcccgggc caagacagat ggacactgcc aagagccgac aatgccctgc tggcctgtct	2007
ccctagggaa ttctgtctat gacagctggc tagcattcct caggaaggac atggggtgcc	2067

c

<210> 2
 <211> 434
 <212> PRT
 <213> Homo sapiens

<400> 2
 Met Glu Val Arg Pro Lys Glu Ser Trp Asn His Ala Asp Phe Val His
 1 5 10 15
 Cys Glu Asp Thr Glu Ser Val Pro Gly Lys Pro Ser Val Asn Ala Asp
 20 25 30
 Glu Glu Val Gly Gly Pro Gln Ile Cys Arg Val Cys Gly Asp Lys Ala
 35 40 45
 Thr Gly Tyr His Phe Asn Val Met Thr Cys Glu Gly Cys Lys Gly Phe
 50 55 60
 Phe Arg Arg Ala Met Lys Arg Asn Ala Arg Leu Arg Cys Pro Phe Arg
 65 70 75 80
 Lys Gly Ala Cys Glu Ile Thr Arg Lys Thr Arg Arg Gln Cys Gln Ala
 85 90 95
 Cys Arg Leu Arg Lys Cys Leu Glu Ser Gly Met Lys Lys Glu Met Ile
 100 105 110
 Met Ser Asp Glu Ala Val Glu Glu Arg Arg Ala Leu Ile Lys Arg Lys
 115 120 125
 Lys Ser Glu Arg Thr Gly Thr Gln Pro Leu Gly Val Gln Gly Leu Thr
 130 135 140
 Glu Glu Gln Arg Met Met Ile Arg Glu Leu Met Asp Ala Gln Met Lys
 145 150 155 160
 Thr Phe Asp Thr Thr Phe Ser His Phe Lys Asn Phe Arg Leu Pro Gly
 165 170 175
 Val Leu Ser Ser Gly Cys Glu Leu Pro Glu Ser Leu Gln Ala Pro Ser
 180 185 190
 Arg Glu Glu Ala Ala Lys Trp Ser Gln Val Arg Lys Asp Leu Cys Ser
 195 200 205
 Leu Lys Val Ser Leu Gln Leu Arg Gly Glu Asp Gly Ser Val Trp Asn
 210 215 220
 Tyr Lys Pro Pro Ala Asp Ser Gly Gly Lys Glu Ile Phe Ser Leu Leu
 225 230 235 240
 Pro His Met Ala Asp Met Ser Thr Tyr Met Phe Lys Gly Ile Ile Ser
 245 250 255

Q1
 Out

Phe Ala Lys Val Ile Ser Tyr Phe Arg Asp Leu Pro Ile Glu Asp Gln
 260 265 270
 Ile Ser Leu Leu Lys Gly Ala Ala Phe Glu Leu Cys Gln Leu Arg Phe
 275 280 285
 Asn Thr Val Phe Asn Ala Glu Thr Gly Thr Trp Glu Cys Gly Arg Leu
 290 295 300
 Ser Tyr Cys Leu Glu Asp Thr Ala Gly Gly Phe Gln Gln Leu Leu Leu
 305 310 315 320
 Glu Pro Met Leu Lys Phe His Tyr Met Leu Lys Lys Leu Gln Leu His
 325 330 335
 Glu Glu Glu Tyr Val Leu Met Gln Ala Ile Ser Leu Phe Ser Pro Asp
 340 345 350
 Arg Pro Gly Val Leu Gln His Arg Val Val Asp Gln Leu Gln Glu Gln
 355 360 365
 Phe Ala Ile Thr Leu Lys Ser Tyr Ile Glu Cys Asn Arg Pro Gln Pro
 370 375 380
 Ala His Arg Phe Leu Phe Leu Lys Ile Met Ala Met Leu Thr Glu Leu
 385 390 395 400
 Arg Ser Ile Asn Ala Gln His Thr Gln Arg Leu Leu Arg Ile Gln Asp
 405 410 415
 Ile His Pro Phe Ala Thr Pro Leu Met Gln Glu Leu Phe Gly Ile Thr
 420 425 430

Gly Ser

<210> 3
 <211> 25
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Putative SXR
 response element from the steroid hydroxylase,
 rCYP3A1

<400> 3
 tagacagttc atgaagttca tctac

25

<210> 4
 <211> 25
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Putative SXR
 response element from the steroid hydroxylase,

rCYP3A2

<400> 4
taagcagttc ataaagttca tctac 25

<210> 5
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Putative SXR
response element from the steroid hydroxylase,
rUGT1A6

<400> 5
actgtagttc ataaagttca catgg 25

<210> 6
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Putative SXR
response element from the steroid hydroxylase,
rbCYP2C1

<400> 6
caatcagttc aacagggttc accaat 26

<210> 7
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Putative SXR
response element from the steroid hydroxylase,
rP450R

<400> 7
cacaggtgag ctgaggccag cagcaggtcg aaa 33

<210> 8
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Putative SXR
response element from the steroid hydroxylase,
rCYP2A1

<400> 8
gtgcaggttc aactggaggt caacatg 27

<210> 9
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Putative SXR
response element from the steroid hydroxylase,
rCYP2A2

<400> 9
gtgctgggttc aactggaggt cagtatg 27

<210> 10
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Putative SXR
response element from the steroid hydroxylase,
rCYP2C6

<400> 10
agtctagttc agtggggggt cagtctt 27

<210> 11
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Putative SXR
response element from the steroid hydroxylase,
hCYP2E1

<400> 11
gagatgggttc aaggaagggt cattaac 27

<210> 12
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Direct repeat
with spacer of 0 nucleotides

<400> 12
catagtcagg tcaaggtcag atcaac 26

<210> 13
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Direct repeat
with spacer of 1 nucleotides

<400> 13
catagtcagg tcataggtca gatcaac

27

<210> 14
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Direct repeat
with spacer of 2 nucleotides

<400> 14
catagtcagg tcaataggtc agatcaac

28

<210> 15
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Direct repeat
with spacer of 3 nucleotides

<400> 15
catagtcagg tcatataggt cagatcaac

29

<210> 16
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Direct repeat
with spacer of 4 nucleotides

<400> 16
catagtcagg tcatataagg tcagatcaac

30

<210> 17
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Direct repeat
 with spacer of 5 nucleotides

 <400> 17
 catagtcagg tcatatatag gtcagatcaa c 31

 <210> 18
 <211> 33
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Description of Artificial Sequence: Direct repeat
 with spacer of 6 nucleotides

 <400> 18
 catagtcagg tcatatataa ggtcaagatc aac 33

 <210> 19
 <211> 33
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Description of Artificial Sequence: Direct repeat
 with spacer of 7 nucleotides

 <400> 19
 catagtcagg tcatatatat aggtcagatc aac 33

 <210> 20
 <211> 36
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Description of Artificial Sequence: Direct repeat
 with spacer of 10 nucleotides

 <400> 20
 catagtcagg tcatatatat ataaggtcag atcaac 36

 <210> 21
 <211> 41
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Description of Artificial Sequence: Direct repeat
 with spacer of 15 nucleotides

 <400> 21
 catagtcagg tcatagtagt agtagtagag gtcagatcaa c 41

<210> 22
 <211> 17
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Example of a
 response element suitable for practice of the
 invention method

<220>
 <221> modified_base
 <222> (7)..(11)
 <223> This region may encompass 5, 4 or 3 nucleotides,
 independently selected from a, c, t or g

<400> 22
 agttcannnn ntgaact

17

<210> 23
 <211> 18
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Example of a
 response element suitable for practice of the
 invention method

<220>
 <221> modified_base
 <222> (7)..(12)
 <223> a, c, t or g

<400> 23
 tgaactnnnn nnaggtca

18

<210> 24
 <211> 18
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic
 oligonucleotide

<400> 24
 tgaactcaaa ggaggtca

18

<210> 25
 <211> 18
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Inverted
 repeat response element with spacer of 0
 nucleotides

<400> 25
 agcttaggtc atgaccta 18

<210> 26
 <211> 19
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Inverted
 repeat response element with spacer of 1
 nucleotides

<400> 26
 agcttaggtc agtgaccta 19

<210> 27
 <211> 20
 <212> DNA
 <213> Artificial Sequence

Q1
 And
 <220>
 <223> Description of Artificial Sequence: Inverted
 repeat response element with spacer of 2
 nucleotides

<400> 27
 agcttaggtc acgtgaccta 20

<210> 28
 <211> 21
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Inverted
 repeat response element with spacer of 3
 nucleotides

<400> 28
 agcttaggtc acagtaccta a 21

<210> 29
 <211> 22
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Inverted
 repeat response element with spacer of 4
 nucleotides

<400> 29
 agcttaggtc acatgtgacc ta 22

<210> 30
 <211> 23
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Inverted
 repeat response element with spacer of 5
 nucleotides

<400> 30
 agcttaggtc acactgtgac cta 23

<210> 31
 <211> 23
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Inverted
 repeat response element with spacer of 6
 nucleotides

<400> 31
 agctttgaac tcaaaggagg tca 23

<210> 32
 <211> 18
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: IR-M

<400> 32
 agcttacgtc atgacgta 18

<210> 33
 <211> 33
 <212> DNA
 <213> Homo sapiens

<400> 33
 tagaatatga actcaaagga ggtagtgag tgg 33

<210> 34
 <211> 33
 <212> DNA
 <213> Homo sapiens

<400> 34
 tagaatatga actcaaagga ggtaagcaaa ggg 33

<210> 35
 <211> 32
 <212> DNA
 <213> Homo sapiens

<400> 35
 tagaatatta actcaatgga ggcagtgagt gg 32

<210> 36
 <211> 25
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic
 oligonucleotide for PCR

<400> 36
 gagcaattcg ccattactct gaagt 25

<210> 37
 <211> 25
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic
 oligonucleotide for PCR

<400> 37
 gtccttgagg tcttctacct ttctc 25

<210> 38
 <211> 25
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic
 oligonucleotide for PCR

<400> 38
 gacgatttgg atctggacat gttgg 25

at
 out

<210> 39
 <211> 15
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic
 oligonucleotide for PCR

<400> 39
 tgaacttcac gaact

15

<210> 40
 <211> 25
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic
 oligonucleotide

<400> 40
 gttttcatct gagcgtccat cagct

25

<210> 41
 <211> 6
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Illustrative
 peptide

<400> 41
 Arg Gly Lys Thr Cys Ala
 1 5

<210> 42
 <211> 15
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic
 oligonucleotide

<400> 42
 tggttcttcac gttct

15

<210> 43
 <211> 15
 <212> DNA
 <213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 43
acaacttcac gaact

15

<210> 44
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Example of a
response element suitable for practice of the
invention method

Q1
Ant
<220>
<221> modified_base
<222> (7)..(11)
<223> This region may encompass 5, 4 or 3 nucleotides,
independently selected from a, c, t or g

<400> 44
aggtcannnn naggtca

17